nature research

Corresponding author(s): Hao Wu

Last updated by author(s): Jul 13, 2023

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	FastQC (v0.11.5), Cutadapt (v1.14), Bismark (v0.18.2), samtools (v0.1.19), Picard (v2.5.0), Drop-Seq Tools (v1.12), STAR (v2.5.2a), DropletUtils (v1.2.2), Scrublet (v0.2.1), IGV (v2.3.91), R (v4.1.1), Seurat (v3.2.3), tidyr (v1.1.4), ggplot2 (v3.3.6), ggtern (v3.3.5), cowplot (v1.1.1), pheatmap (v1.0.12), FlowJo (v10.5)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw and processed data reported in this paper are available at the Gene Expression Omnibus (GEO) with accession number GSE236798. All metadata for snhmC-seq and Joint-snhmC-seq experiments are included in Supplementary Table 1,4, 5, and 6. The following publicly available datasets were used in this study: GSE97179 (snmC-seq). The plasmid used for purification of in-house APOBEC3A is available through Addgene (#187822).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. The sample size (snhmC-seq: 360 nuclei, Joint-snhmC-seq: 545 nuclei (paired snhmC-seq2/split: and snmC-seq2/split assays), sNucDrop-seq: 6,000 nuclei) were in the range of other similar studies on single-cell DNA methylome experiments (Luo et al, Nat. Commun., 2018; Nichols et al, Nat. Commun., 2022) or single-cell RNA-seq experiments (Macosko et al, Cell, 2015; Hu et al, Mol. Cell, 2017).
Data exclusions	Data were not excluded.
Replication	Three biologically independent mouse brains were analyzed. We assessed the replication by examination of intermingling of cells originating from different samples in clusters and in low-dimensional UMAP space. All attempts at replication were successful and shown in the manuscript.
Randomization	Randomization was not applied. All cells pass quality filtering are used for the analysis and shown in the manuscript.
Blinding	Not applied. Blinding was not relevant since sample identities were encoded into experiment design, and then subjected to single-cell sequencing analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Wild-type mESC (J1 line, ATCC, SCRC-1010) were originally purchased from ATCC and no additional authentication was performed. TET TKO mESC (J1) line was derived in house by CRISPR/Cas9 genome edited using previously validated sgRNAs from Wang et al, Cell (2013). The Tet-TKO mESC line was confirmed as noted below.
Authentication	Genotype of Tet-TKO mESCs (J1) were verified by both Sanger sequencing and single-cell RNA-seq at Tet1/2/3 loci. The lack of 5-hydroxymethylcytosine (5hmC) was confirmed by mass spectrometry as previously described (Schutsky et al., Nat Biotech 2018).
Mycoplasma contamination	Not tested.
Commonly misidentified lines (See ICLAC register)	None of the cell lines used were listed in ICLAC.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Mice were group-housed in cages of three to five in a 12-hr light/dark cycle (light on at 7:00 am) with food and water provided ad libitum. Consistent relative air humidity of 55% and room temperature of 22°C. All mice used for experiments were naïve to behavioral assays and other procedures. For all bulk-nucleus and single-nucleus experiments, 6-10 weeks old, wild-type, male mice in a pure C57BL/6J background were used. For prospectively isolating inhibitory and excitatory neuronal subtypes to benchmark snhmC-seq and joint-snhmC-seq, 9-week-old transgenic, male mice (Mecp2Tavi/y, NexCre/+, Rosa26BirA/BirA) from a C57BL/6J background were used (Johnson et al, Nature Medicine 2017).
Wild animals	No wild animals were used in the study.
Field-collected samples	No field collected samples were used in the study.
Ethics oversight	Experiments were conducted in accordance with the ethical guidelines of the National Institutes of Health and with the approval of the Institutional Animal Care and Use Committee of the University of Pennsylvania. Animals were solely used for collection of material, and no additional animal experiments were performed.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Mouse brains were rapidly resected on ice. Cortices were flash frozen in liquid nitrogen for 2 min and subsequently kept at -80°C before nuclear isolation. Nuclei were isolated and purified by surcose ultracentrifugation.
Instrument	BD Influx™ cell sorter
Software	FlowJo (v 10.5)
Cell population abundance	The purity was verified by the clustering analysis from single nuclei RNA-seq.
Gating strategy	After gating on forward scatter (FSC) vs. side scatter (SSC), doublets were excluded by FSC-H vs. FSC-W and SSC-H vs. SSC-W gating. Nuclei were identified by DAPI staining. Neuronal and non-neuronal nuclei was gated based on the NeuN antibody, respectively. Streptavidin is conjugated with PE was further used to distinguish the excitatory neuron (PE+/DAPI+) from inhibitory neuron (PE-/DAPI+) nuclei. Representative gating strategy is provided in Extended Data Fig. 4.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.